AD_____

Award Number: W81XWH-06-1-0091

TITLE: Predicting Prostate Cancer Recurrence by Gene Expression Analysis of Formalin-Fixed, Paraffin Embedded Tissue

PRINCIPAL INVESTIGATOR: Ann Schwartz, Ph.D.

CONTRACTING ORGANIZATION: Wayne State University
                              Detroit, MI 48201

REPORT DATE: November 2008

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
                  Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
                              Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1 November 2008	Annual Report	30 OCT 2007 - 29 OCT 2008

4. TITLE AND SUBTITLE

Predicting Prostate Cancer Recurrence by Gene Expression Analysis of Formalin-Fixed, Paraffin Embedded Tissue

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-06-1-0091

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Ann Schwartz, Ph.D., MPH Isaac Powell, M.D.; Wael Sakr, M.D.; Lance Heilbrun, Ph.D.

E-Mail: schwarta@karmanos.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Wayne State University
Detroit, MI 48201

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Prostate cancer is the leading cancer diagnosed in men and accounts for approximately 30,000 deaths per year in the US. In addition to large racial differences in outcome, other factors affecting prognosis include clinical stage, Gleason grade, and PSA levels. The genetic contribution to prostate cancer risk is well accepted, but less has been done to evaluate the genetic contribution to recurrence risk and variation by race. With the advent of newer technologies, discovery of molecular signatures of prognosis are now possible and are the focus of this study. Using Illumina's DASL assay, we are evaluating 529 genes for expression differences comparing men with, and without, recurrence. The gene expression sets for African American men will be compared to those for white men to identify genes contributing to racial disparities in outcome. A unique and diverse patient population has been identified, and 649 tumor samples have been processed for gene expression studies of prostate cancer recurrence and racial disparities in outcome. We finalized the panel of genes to be included in the expression array and production of the custom assay has been completed. All clinical data has been assembled. The DASL assay has been run on our samples and these gene expression results are being analyzed.

15. SUBJECT TERMS
Prostate Cancer, Race, Gene Expression

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	UU	13	**19b. TELEPHONE NUMBER** (include area code)
U	U	U			

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Ann G. Schwartz, PhD, MPH
W81XWH-06-1-0091

Introduction: Prostate cancer is the leading cancer diagnosed in men and accounts for approximately 30,000 deaths per year in the US (1). Large racial disparities in outcome are seen, with African American men having poorer survival after a diagnosis than white men (2). In addition to racial differences in outcome (both recurrence and overall survival), factors affecting prognosis include clinical stage, Gleason grade, and PSA levels (2, 3). The genetic contribution to prostate cancer risk is well accepted, but less has been done to evaluate the genetic contribution to recurrence risk and the variation by race. With the advent of newer technologies, discovery of molecular signatures of prognosis are now possible and are the focus of this study. Using a new application, Illumina's DASL assay, we are evaluating 529 genes for expression differences comparing men with, and without, recurrence. The gene expression sets for African American men will be compared to those for white men to identify genes contributing to racial disparities in outcome.

Body: The study population has been finalized to include men diagnosed with prostate cancer who underwent radical prostatectomies for clinically localized prostate cancer from January 1991 through June 1996. 649 men (275 African American and 374 white) have sufficient follow-up and tumor blocks for inclusion. These men also have data available on: age, clinical stage, preoperative Gleason score, preoperative PSA, preoperative hormonal therapy, post-operative stage, nodal status, Gleason score, capsular, margin or seminal vesicle involvement, tumor volume, treatments, and postoperative PSA. An analytic dataset of these variables has been developed. Long-term survival data and biochemical recurrence data have been reviewed for completeness and steps have been taken to fill in missing data gaps.

The proposed panel of genes has been revised to include newly identified regions associated with prostate cancer recurrence and racial disparities. We worked closely with Illumina to design a custom panel for the DASL assay, and Illumina has completed production of the panel. Quality control tests have been completed on the extracted RNA and the assay has been successfully run on all samples providing the expression data. Data analysis is now underway.

Key Research Accomplishments:
- A unique and diverse patient population has been identified.
- RNA has been isolated from 649 tumor samples.
- Outcome data has been reviewed for completeness, and steps have been taken to fill in missing data gaps.
- An analytic dataset of clinical characteristics for the final patient list has been developed (see detail below).
- The gene expression list was finalized and Illumina has completed the development of the customized assay.

- The DASL assay has been run on all samples and expression data are now being analyzed (see preliminary data analysis below).

Specifically, descriptive analyses were performed on all the variables in the clinical dataset. Variables were compared by race (African American versus White, excluding five Pacific Islanders). Wilcoxon tests were performed on continuous data and Chi-square tests were used for categorical data. When the count was small in the cross-tabulate, a Fisher's exact test was performed. Age and PSA distributions were plotted using density and box-plot within race groups. KM curves of overall survival were plotted for Black and White. A log-rank test was used to compare the overall survival between the two race groups. Since prostate cancer patients often die from causes other than the disease, competing risk analyses were done using the cumulative incidence rates and Gray's test.

The descriptive analyses results are shown in Table 1. There are significant differences between race groups on PSA level, primary Gleason grade, Gleason Score category 2, and overall survival. The longest follow-up time for survival was more than 20 years (Figure 2). When competing risks were accounted for, the risk of death due to prostate cancer was not significantly different between two race groups (Gray's test p value = .33 in Figure 3)

Table 1. Descriptive analyses results

	Race		P-value
	Black (N=275)	White (N=373)	
Median Age	64	62	0.0634†
Median baseline PSA	8.5	7.5	0.0018†
	34 (12%) missing	48 (13%) missing	
Gleason Grade			
Gleason grade primary			
0	3	0	0.0055‡
2	11	12	
3	158	259	
4	101	99	
5	2	2	
Gleason grade secondary			
0	3	0	0.3262‡
1	0	1	
2	18	27	
3	130	174	
4	105	151	
5	19	19	
Gleason score counts			
Category 1			0.4118§
6 or less	92	141	
7	140	184	

8 or more	43	48	
Category 2			$0.0273^§$
7 (3+4) or less	172	264	
7 (4+3) or greater	103	109	
Pathology Stage, T value from TNM			$0.9118^§$
Stage II	178	243	
Stage III	97	130	
Median overall survival in years (95% CI)	12.9 (12.5, 13.3) 7 (3%) missing	14.2 (13.8, 14.3) 7 (2%) missing	$<0.0001^{§§}$

Notes : †: Wilcoxon test, §: chi-square test, §§: log-rank test ‡: Fisher's exact test

Figure 1. Age and PSA distributions

PSA Distributions

PSA Distributions
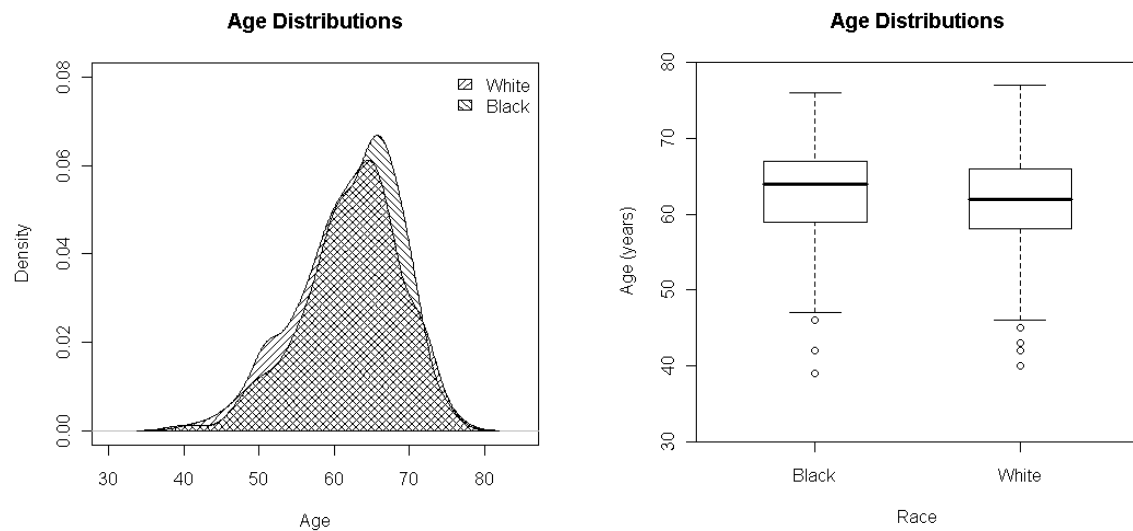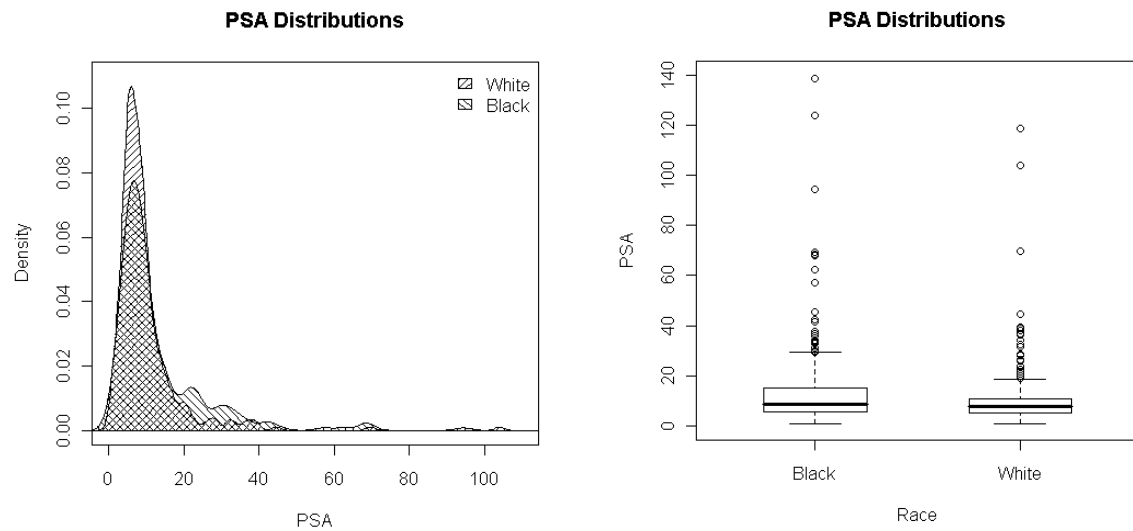
Figure 2. KM Curves of Overall Survival for Black and White
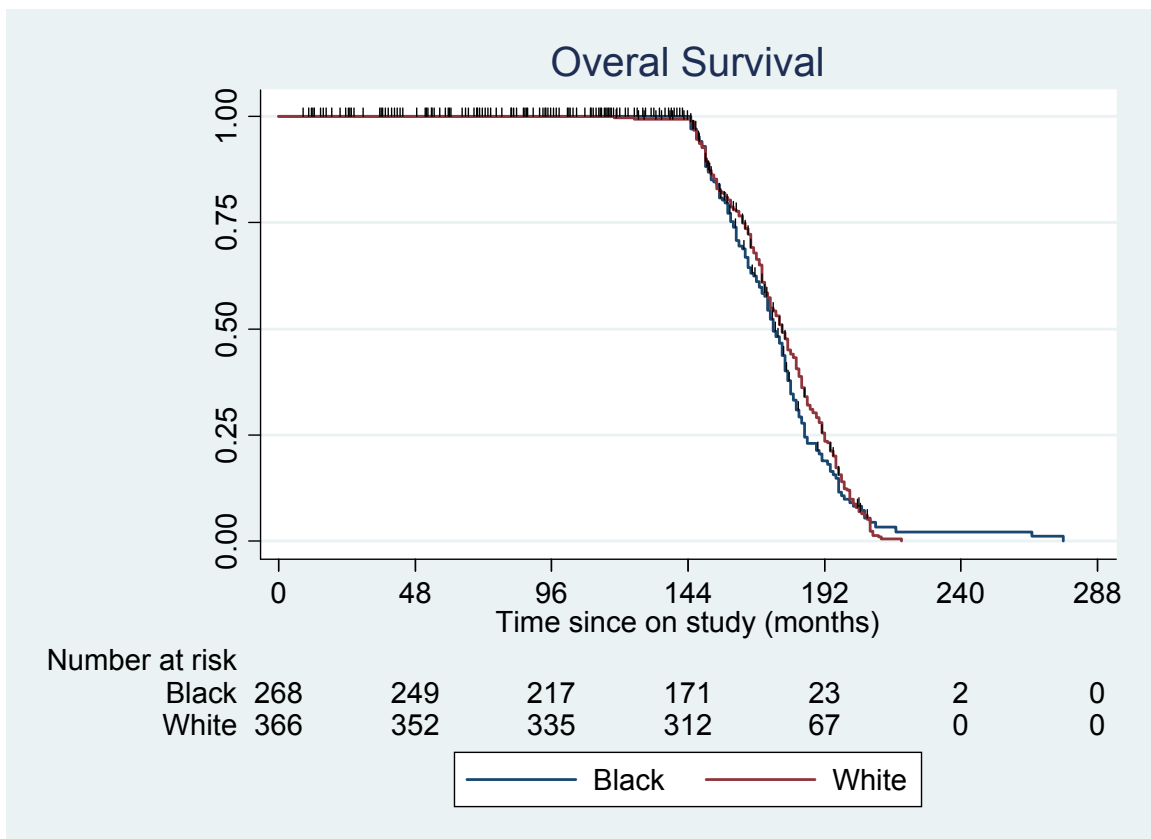
Figure 3. Cumulative Incidence Functions and Gray's test

Prostate cancer: CIFs for disease failure

Prostate cancer: CIFs for competing risk

(a) Death with disease failure

(b) Death without disease failure
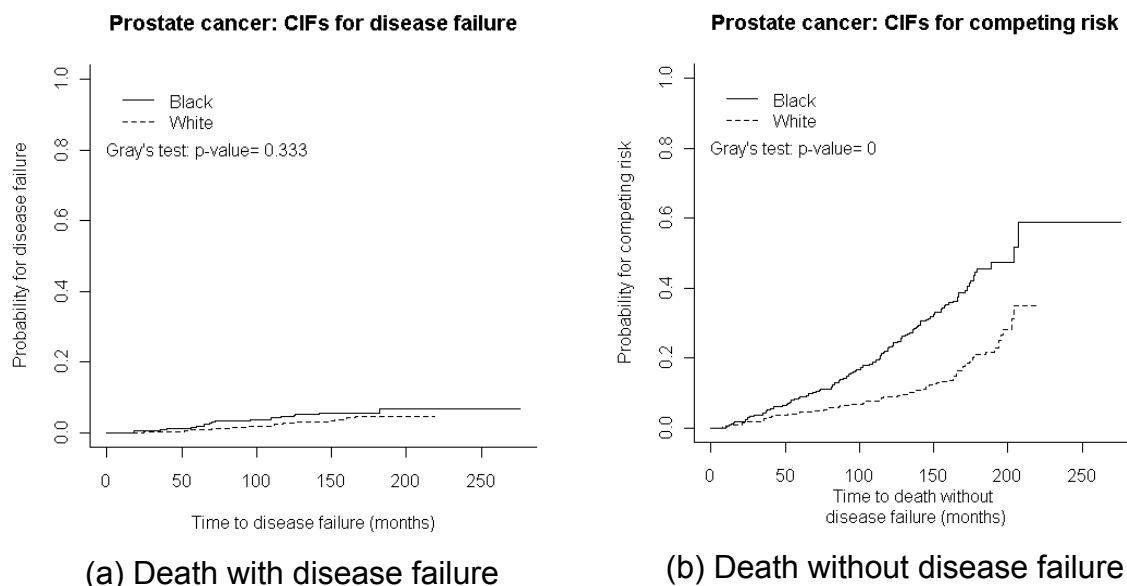
656 prostate FFPE samples were extracted for RNA and prepared for microarray analysis of gene expression by the Illumina DASL protocol.

A custom array of 529 prostate cancer related genes was developed and interrogated with an average of 3 probes per gene (1536 probes total) in the DASL bead pool protocol, using the Illumina Sentrix chip format (96 individual sample arrays).

The 656 samples + 16 duplicates (= 672 total samples) were run on 7 Sentrix chips (named AS1 – AS7). Each Sentrix chip contained 96 individual sample arrays (7 x 96 = 672 total samples). Each Sentirx chip contained the Illumina negative controls for background and noise estimation and positive controls to monitor probe hybridization and array processing protocols.

Initial analysis of microarray data:
A. Procedure: Sentrix chips were hybridized and read in the Illumina Bead Scanner. Bead scanner image data files for all of the 672 individual arrays (7 Sentrix chips x 96 arrays / chip) were imported into Illumina Bead Studio v.3.4 software for gene expression analysis. Each Sentrix chip (96 arrays) was maintained as a separate analysis group, in order to enable quality control analysis and assessment of technical variability in array processing. Microarray data were quantile normalized and scaled for comparison among individual arrays.

B. Results of initial analysis

Quality control:

1. The average signal intensity distribution was determined for the set of 672 arrays (Figure 4), and compared to the background determined from the internal negative controls on each array (Figure 5). These results demonstrate a clear separation of signal (mean = 7489 units, cv = 0.14) above the background (mean = 708, cv = 0.21).

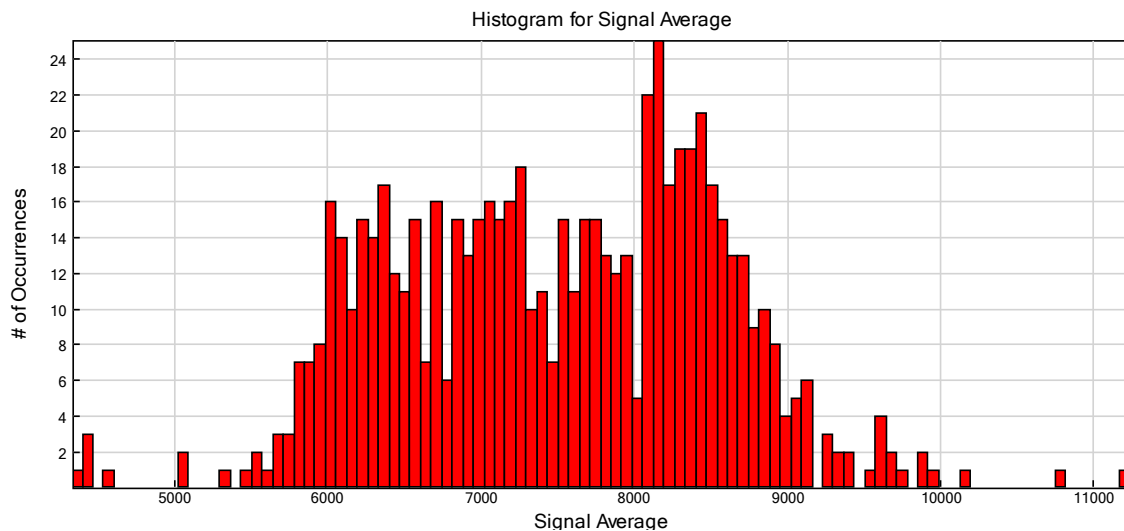Figure 4: Histogram of average signal intensity for all 672 normalized array samples
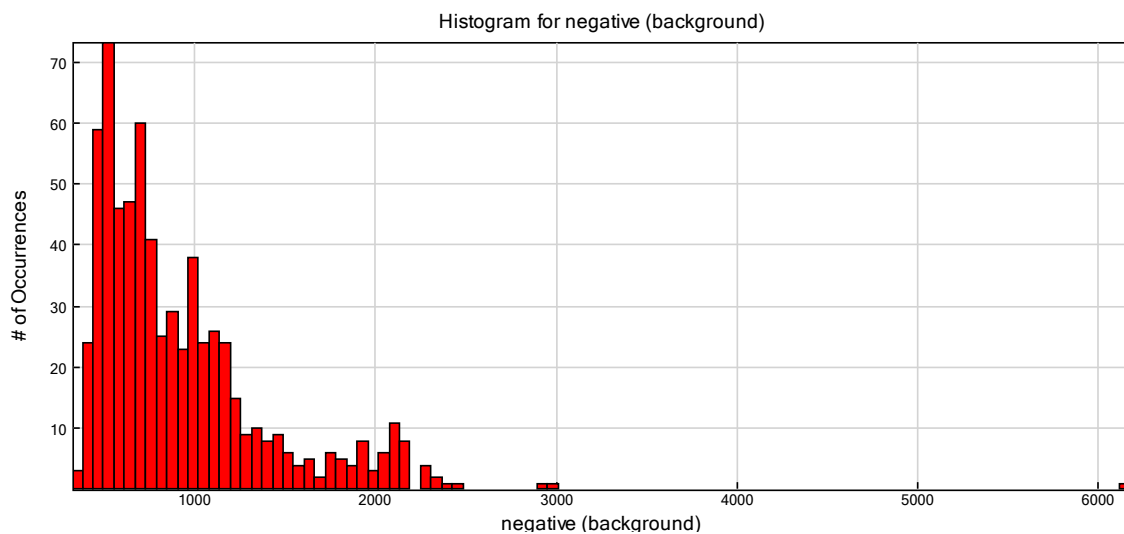

Figure 5: Histogram for the background (internal negative controls) for all 672 arrays

It is possible to estimate the "noise" on each array, as distinct from the total background, using Illumina internal controls. This distribution is given in Figure 6, and "looks better" than the background distribution in Figure 6 (mean = 692, cv = 0.19).

Figure 6: Histogram of the calculated "noise" (from internal negative controls) for all 672 arrays
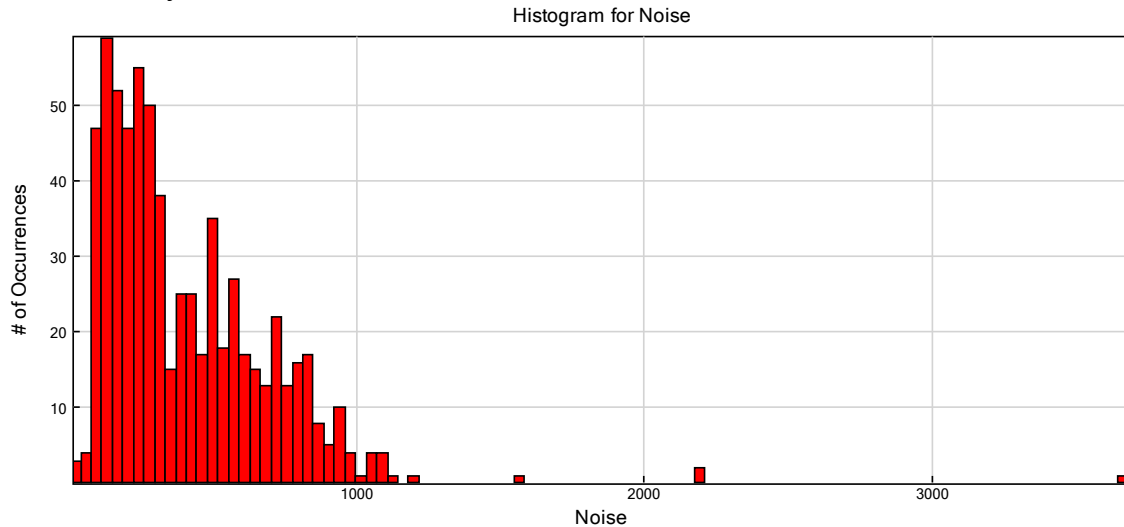
2. The presence of each of the 529 target genes was determined for each of the individual arrays (p = 0.05 level) and the distribution plotted (Figure 7), which demonstrates that the vast majority of the candidate prostate genes were in fact expressed and detected in the samples.

Figure 7: Histogram for the detection of the 592 target genes on the custom DASL array
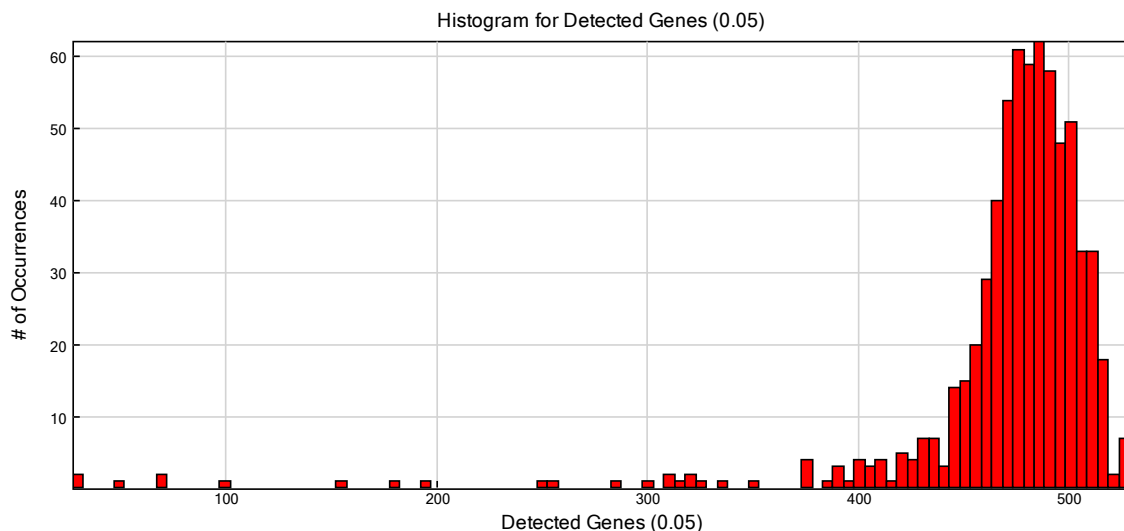
3. The average expression level of each probe (2 – 3 probes / target gene) was compared across the Sentrix chips and plotted as a heat map (Figure 8, approximately 20 probes, grouped as the 2 – 3 genes associated with each target gene are presented) against a clustering (Manhattan distance, k-means algorithm) of the Sentrix chips. This heat map indicates that expression values for individual probes and genes are consistent across the 7 Sentrix chips. Sentrix chip # AS3 has an average expression level that is lower than the other chips due to a large number of failed RNA probe preparations for this set of samples. These "failed" outliers will be removed for subsequent comparative analysis.

Figure 8: Heat map showing Sentrix chip (AS1 – AS7) clustering against the individual probes (GI_nnnnnn) and target genes on the 672 arrays (Note: a selection of ~ 20 representative probes of the 1536 total probes on the custom array)
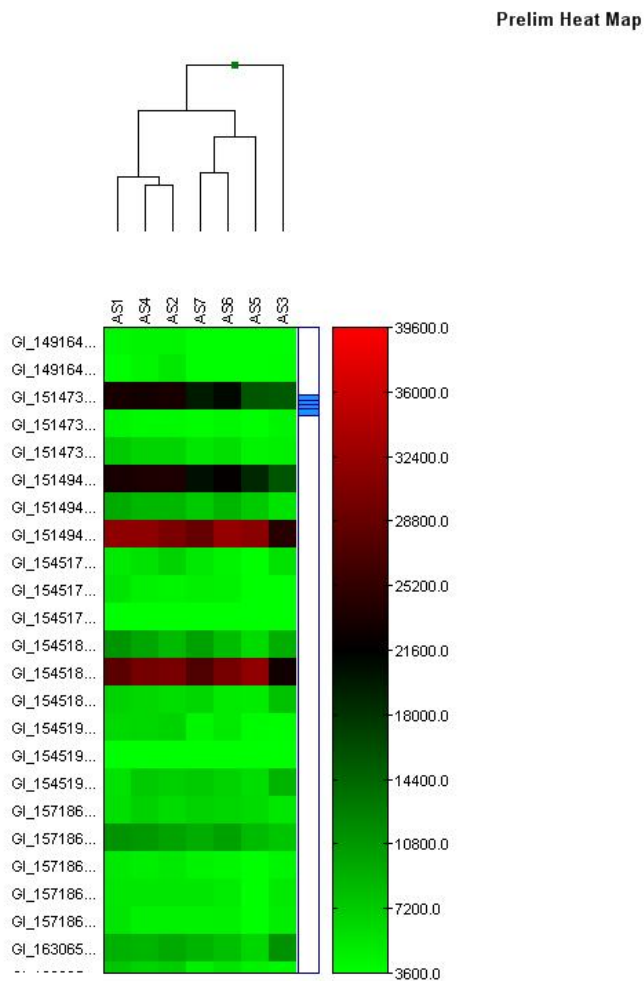
4. The average signal intensity distribution for each of the 7 Sentrix chips was visualized as a box plot in Figure 9. These data demonstrate that the quantile normalization effectively scales the different chips. The removal of "failed" outliers will further improve the normalization across samples and chips.

Figure 9: Box plot of the average signal intensity distribution for each of the 7 Sentrix chips (AS1 – AS7) (Note: There are 1536 probes total for the 529 target genes.)
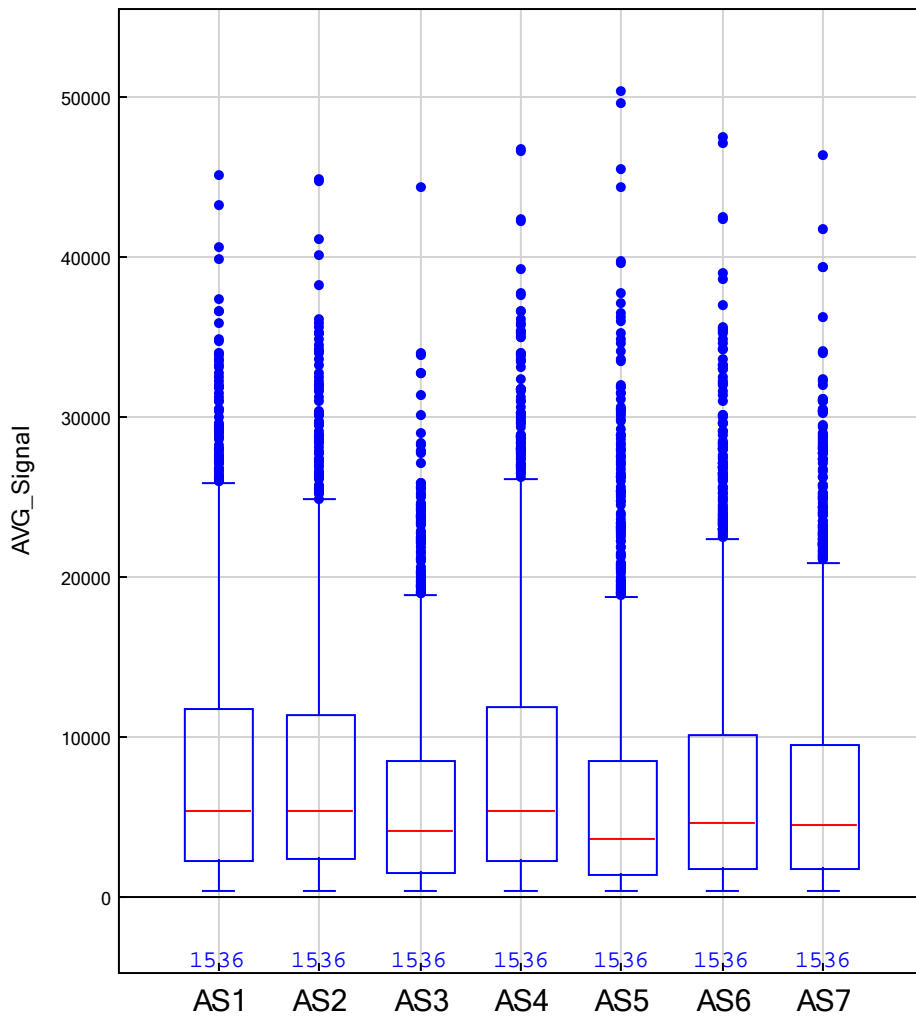
Conclusions from the initial microarray analysis:
1. The DASL procedure is largely effective in extracting RNA for subsequent microarray analysis from FFPE prostate samples.

2. The set of 529 target genes in the custom prostate specific micorarray panel are well expressed (signal to background values) in the prostate samples, and are well detected by the custom DASL array.

3. Quality control metrics demonstrate that samples processed on different Sentrix chips can be normalized and are appropriate for subsequent group wise comparative analysis.

Reportable Outcomes: We anticipate the completion of the analysis by the end of the grant period with manuscript preparation to follow.

Conclusion: We are in the final stage of this project, namely data analysis and manuscript preparation and dissemination. The resulting gene expression sets for African American men and white men will lead to the identification of genes contributing to racial disparities in outcome after a prostate cancer diagnosis, leading to a better understanding of the carcinogenic process and the potential development of novel therapies.

References:
1. Jemal A, Murray T, Samuels A, et al. Cancer Statistics, 2003. CA Cancer J Clin 2003; 53:5-26.
2. Powell IJ, Heilbrun LK, Sakr W, et al. The predictive value of race as a clinical prognostic factor among patients with clinically localized prostate cancer: a multivariate analysis of positive surgical margins. Urology 1997; 49:726-31.
3. Partin AW, Kattan MW, Subong EN, et al. Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional update. JAMA 1997; 277:1445-1451.

Appendices: NA